

De l'analyse d'un corpus de texte à la conception d'une interface graphique facilitant l'accès aux connaissances sur le médicament

Jean-Baptiste Lamy, Catherine Duclos, Alain Venot

Laboratoire d'Informatique Médicale et de Bioinformatique (LIM&BIO), UFR SMBH, Université Paris 13, 74 rue Marcel Cachin, 93017 Bobigny cedex, France
jiba@soyaproject.org, catherine.duclos@avc.ap-hop-paris.fr,
avenot@smbh.univ-paris13.fr

Résumé :

Pour faciliter l'accès aux contre-indications, interactions médicamenteuses et effets indésirables des médicaments par les professionnels de santé, nous avons conçu une interface graphique s'appuyant sur un langage iconique. Dans cet article, nous présentons la méthode de conception de cette interface, laquelle repose sur des techniques de visualisation d'information, et sur l'analyse distributionnelle d'un corpus de textes de référence sur le médicament. Cette analyse a permis de déterminer les principaux axes pour organiser les icônes sur l'interface.

Mots-clés : Interface graphique, Interface iconique, Accès aux connaissances textuelles, Analyse d'un corpus de texte, Connaissances sur le médicament.

1 Introduction

Les erreurs de prescriptions sont responsables d'un nombre important d'hospitalisations (Moore *et al.*, 2007) ; mais beaucoup d'entre elles auraient pu être évitées si les propriétés des médicaments avaient été prises en compte lors de la prescription. Ces connaissances sont disponibles sous forme textuelle dans le RCP (Résumé des Caractéristiques Produits, le texte de référence sur un médicament décrivant notamment contre-indications, interactions médicamenteuses et effets indésirables). Cependant, le volume de texte et le manque de temps empêchent souvent la lecture du RCP en consultation.

Les capacités de la vision humaine (Paivio, 1990) étant sous-utilisées par la présentation textuelle, l'accès aux connaissances peut être facilité par une présentation graphique. Les approches graphiques peuvent être classées en deux catégories : les *langages iconiques* (Chang, 1986) qui permettent de représenter

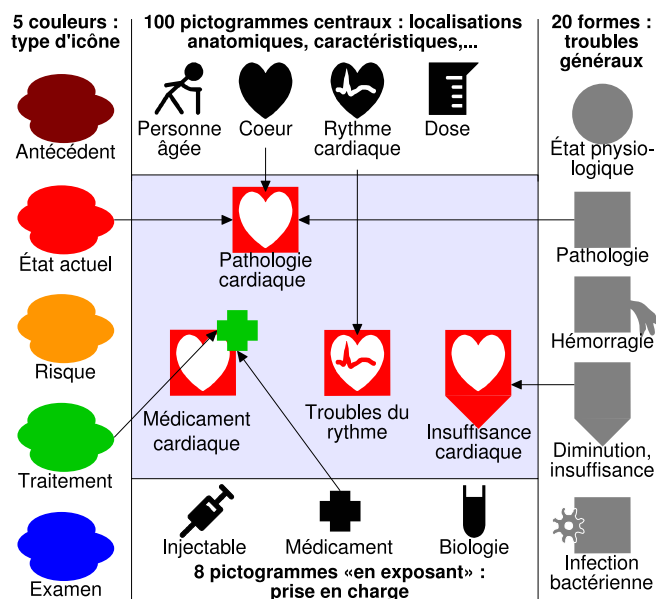


FIG. 1 – Construction des icônes du langage VCM à partir des primitives.

des informations ou des connaissances en combinant pictogrammes, couleurs, formes,..., selon une grammaire graphique (Meunier, 1998), par exemple les panneaux routiers, et la *visualisation d'information* (Andrews, 2002) qui propose plusieurs techniques pour présenter graphiquement la structure de l'information, par exemple l'arborescence d'un système de fichiers. Bien que complémentaires, ces deux approches ont rarement été utilisées de manière conjointe.

Pour faciliter l'accès aux connaissances médicales, nous avons mis au point le langage iconique VCM (Visualisation des Connaissances Médicales) (Lamy *et al.*, 2008a; Lamy *et al.*, 2008b) qui permet de représenter par des icônes les principaux concepts des connaissances médicales : pathologies, risques, médicaments, examens, habitudes de vie,... Il se compose d'un jeu de primitives graphiques incluant plusieurs couleurs et une centaine de pictogrammes, et d'une grammaire graphique décrivant les règles pour combiner les primitives et construire les icônes. Par combinatoire, il est ainsi possible de créer plusieurs milliers d'icônes (figure 1). La grammaire du langage VCM définit aussi des relations *est-un* entre les icônes : par exemple entre l'icône signifiant "trouble du rythme" et celle signifiant "pathologie cardiaque". VCM a été conçu à destination des professionnels de santé (médecins, pharmaciens...). Il demande un court apprentissage, qui est facilité par l'utilisation de nombreux symboles et pictogrammes évocateurs.

Nous nous intéressons dans cet article à la conception d'une interface graphique pour faciliter et accélérer la consultation des contre-indications, des interactions médicamenteuses et des effets indésirables décrits dans un RCP, en s'appuyant sur le langage VCM et sur des techniques de visualisation d'information. Nous

avons travaillé sur deux cas d'utilisation mis au point par un expert du domaine :

1) le médecin prescrit un médicament complexe ou qu'il connaît mal, et souhaite vérifier l'ensemble des contre-indications et interactions médicamenteuses par rapport à son patient. Il a alors besoin d'une vue d'ensemble des propriétés du médicament, ce que ne fournit pas le texte.

2) le patient présente une pathologie ou un traitement inhabituel, et le médecin veut vérifier si le médicament qu'il prescrit est contre-indiqué avec cette pathologie ou ce traitement ; ou alors le patient présente un symptôme et le médecin veut vérifier s'il s'agit d'un effet indésirable d'un médicament donné. Si la contre-indication, l'interaction médicamenteuse ou l'effet indésirable recherché n'existe pas, cette connaissance est implicite dans le texte du RCP ; par exemple si le RCP ne mentionne pas de contre-indication avec l'asthme, il est sous-entendu qu'il n'y en a pas. Le médecin doit alors parcourir la totalité de la section correspondante du RCP avant de pouvoir déduire son absence, ce qui prend du temps. Un second besoin du médecin est donc d'accéder rapidement aux absences de contre-indication, d'interaction médicamenteuse ou d'effet indésirable.

L'objectif de cet article est de présenter la méthode de conception d'une interface répondant à ces cas d'utilisation. Tout d'abord, nous présenterons une analyse d'un corpus de RCP qui nous a permis de choisir les axes pour classer et organiser les contre-indications, les effets indésirables et les interactions médicamenteuses sur une vue d'ensemble. Ensuite, nous verrons comment les résultats de cette étude ont été utilisés pour construire une interface graphique s'appuyant sur des techniques de visualisation d'information. Enfin, nous donnerons brièvement les résultats obtenus lors d'une évaluation de cette interface, et nous discuterons la méthode utilisée.

2 Analyse d'un corpus de RCP

2.1 Matériel et méthodes

Nous avons travaillé sur les 278 RCP des médicaments disponibles en France et présents dans la liste des médicaments essentiels de l'OMS, donnant lieu à 3 corpus comprenant les sections contre-indications (constituées d'énumérations et comprenant 25672 mots), effets indésirables (constituées d'énumérations et de phrases, et comprenant 534376 mots) et interactions médicamenteuses (constituées d'énumérations et comprenant 492351 mots) de chaque RCP.

Cette analyse distributionnelle s'est appuyée sur des outils de Traitement Automatique de la Langue (Bouillon & Vandooren, 1998). Elle a consisté à extraire les candidats termes automatiquement avec Cordial et Lexter (Bourigault, 1995), puis à extraire les termes décrivant des contre-indications pour le premier corpus, des effets indésirables pour le second, et des médicaments interagissant pour le troisième, et enfin à extraire les attributs qui caractérisent chacun de ces termes, de manière semi-automatique, selon les étapes suivantes (figure 2) :

(1) Nous appellerons CT1 la liste des candidats-termes. À partir de CT1, extraire la liste de tous les noms et adjectifs, sans doublon, que nous appellerons W1.

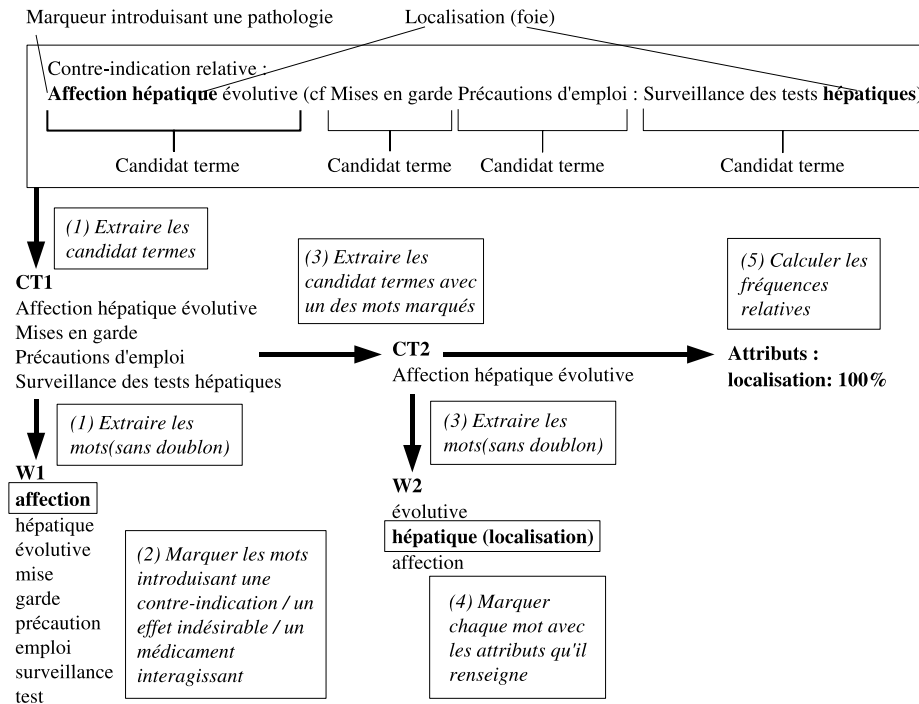


FIG. 2 – Méthode utilisée pour l'analyse des corpus de textes. La méthode est ici appliquée à une contre-indication.

(2) Un expert marque dans la liste W1 quels sont les mots indiquant les pathologies contre-indiquées pour le premier corpus (par exemple “maladie”, “cardiopathie”, “hypertension”,... indiquent des pathologies), les effets indésirables pour le second, et les médicaments interagissant pour le troisième.

(3) À partir de CT1, extraire la liste de tous les candidats termes comprenant au moins un mot marqué à l'étape précédente dans W1 ; nous l'appellerons CT2. Cette liste contient donc tous les termes qui nous intéressent. Ensuite, à partir de CT2, extraire la liste de tous les noms et adjectifs, sans doublon, que nous appellerons W2.

(4) Un expert marque dans la liste W2 chaque mot avec les attributs qu'il renseigne, y compris les attributs qui ne sont pas indiqués de manière explicite mais dont un médecin a connaissance. Par exemple, “cardiopathie” renseigne l'attribut “localisation anatomo-fonctionnelle” (valeur : cardiaque), et “tuberculose” renseigne les attributs “localisation anatomo-fonctionnelle” (valeur : pulmonaire) et “étiologie” (valeur : bactérienne).

(5) Pour chaque attribut, la fréquence relative de l'attribut est donnée par la proportion de candidat-termes dans la liste CT2 qui inclut au moins un mot marqué avec cet attribut dans la liste W2.

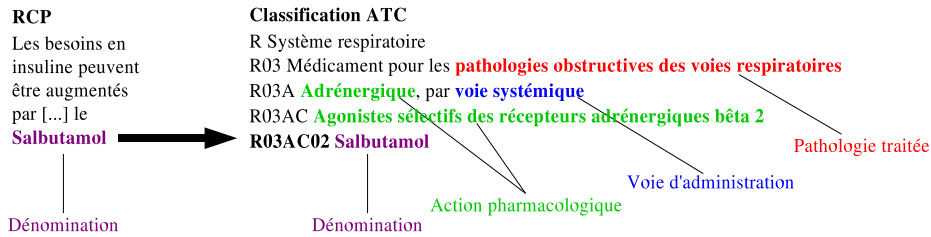


FIG. 3 – Exemple d'utilisation de la classification ATC pour caractériser les dénominations des médicaments. Ici, la dénomination chimique "Salbutamol" est enrichie par des attributs concernant l'activité pharmacologique, la voie d'administration et la pathologie traitée.

Les étapes 2 et 4 ont été réalisées à la main par un expert du domaine, pharmacien de formation. Les autres étapes ont été automatisées à l'aide de scripts.

Pour les médicaments, les dénominations chimiques des principes actifs sont fréquemment employées, et il est difficile, même pour un expert, de connaître la totalité des attributs se rapportant à chaque dénomination. Nous avons résolu ce problème avec la classification ATC (*Anatomical Therapeutic Chemical drug classification*, classification Anatomique, Thérapeutique et Chimique des médicaments). Dans un premier temps, les libellés des termes ATC ont été marqués avec les attributs qu'ils définissaient. Dans un second temps, nous avons relié chaque dénomination de médicament au terme ATC correspondant, et nous lui avons associé les attributs marqués dans le libellé, ainsi que dans ceux de tous les termes parents dans la classification (figure 3).

2.2 Résultats

L'analyse du corpus de RCP a permis de déterminer les fréquences relatives des attributs caractérisant les contre-indications et les effets indésirables (tableau 1) et les médicaments (tableau 2).

Pour les contre-indications et les effets indésirables, une localisation anatomo-fonctionnelle (par exemple cardiaque ou pulmonaire) peut être associée à 69% des contre-indications et 76% des effets indésirables, un trouble (c'est à dire le problème présent, par exemple une insuffisance dans la fonction d'un organe) à 46% des contre-indications et 70% des effets indésirables, et une étiologie (c'est à dire la cause de la pathologie, par exemple virale ou bactérienne) à 30% des contre-indications et 10% des effets indésirables. Les autres attributs, gravité de la pathologie, niveau de certitude du diagnostic, et nom propre associé, sont plus rarement présents.

Pour les médicaments, les dénominations (chimiques ou noms de marque) ont été remplacées par les attributs correspondant en s'appuyant sur la classification ATC. Un terme ATC a pu être trouvé pour 86% des dénominations, les autres correspondant à des familles chimiques très larges ou sans équivalent ATC. Cette

	Contre-indications	Effets indésirables	Pathologies traitées par les médicaments
Localisation anatomo-fonctionnelle	69%	76%	39%
Troubles	46%	70%	34%
Étiologie	30%	10%	31%
Gravité	10%	5%	-
Incertitude	2%	0,2%	-
Nom propre	2%	2%	-
Localisation et / ou étiologie	94%	82%	67%
Localisation et / ou trouble	73%	84%	40%

TAB. 1 – Fréquences relatives d'utilisation des attributs pour les pathologies et effets indésirables.

	Médicaments
Dénomination chimique	71%
Pathologie traitée	67%
Effet	26%
Voie d'administration	26%

TAB. 2 – Fréquences relatives d'utilisation des attributs pour les médicaments.

correspondance a été réalisée de façon automatique pour 56% des termes (qui correspondent aux dénominations chimiques bien orthographiées), et de façon manuelle pour le reste (noms de marque, familles chimiques ou dénominations chimiques comportant une faute d'orthographe). L'activité des médicaments est caractérisée par la pathologie qu'il traite (par exemple "anti-hypertenseur") dans 67% des cas, et par l'effet (opposé à la pathologie traitée, par exemple "hypotenseur") dans 26%. Pour les médicaments décrits par la pathologie traitée, nous avons indiqué dans le tableau 1 les attributs caractérisant la pathologie traitée.

L'utilisation conjointe des attributs localisation anatomo-fonctionnelle et étiologie permet de couvrir un maximum de termes rencontrés dans les RCP : pour 94% des contre-indications, 82% des effets indésirables et 67% des médicaments il est possible d'associer une localisation ou une étiologie (contre 73%, 84% et 40% si l'on retenait les attributs localisation et trouble). Les termes associés ni à une localisation ni à une étiologie sont principalement des candidats termes non interprétables médicalement (par exemple des démonstratifs : "cette pathologie") ou mal extraits par l'analyseur syntaxique, ou, pour les médicaments, des classes pharmacologiques ou chimiques très larges ("les sulfamides").

En conclusion de cette analyse, nous avons retenu les deux axes localisation anatomo-fonctionnelle et étiologie pour organiser les contre-indications, les effets indésirables et les interactions médicamenteuses sur l'interface interactive.

3 Conception de l'interface

3.1 Matériel et méthodes

Afin de présenter une vue d'ensemble des contre-indications, interactions médicamenteuses et effets indésirables d'un RCP, nous avons utilisé une technique de visualisation d'information appelée *Fisheye* (Furnas, 1986; Hascoët & Beaudouin-Lafon, 2001). Cette technique consiste à séparer les informations à visualiser en deux parties : le *contexte*, qui donne une vue d'ensemble de la totalité des informations avec un faible niveau de détails, et le *focus*, qui présente en détail la partie des informations qui a été sélectionnée par l'utilisateur dans le contexte. Dans notre interface, le contexte sera l'ensemble des icônes VCM associées aux contre-indications, effets indésirables et interactions médicamenteuses du RCP, et le focus, les passages textuels du RCP associés aux icônes sélectionnées par l'utilisateur.

L'analyse du corpus de RCP a conduit à organiser le contexte selon deux axes anatomique et étiologique. Un schéma anatomique peut être soit être réaliste, soit simplifié et organisé en "cases" fixes et semi-arbitraires, correspondant chacune à une localisation (figure 4). Nous avons préféré le schéma simplifié pour 3 raisons : (1) sur un schéma réaliste, il est difficile de représenter simultanément les organes de certaines régions comme l'abdomen à cause de leur superposition, (2) les cases ont toutes la même taille ce qui évite des erreurs d'interprétations, par exemple d'accorder plus d'importance aux organes de plus grande taille, et (3) il est plus facile de placer des icônes VCM sur un schéma simplifié. Le schéma anatomique a ensuite été enrichi en ajoutant des cases pour les étiologies ; ces cases ont été placées à l'extérieur du personnage, car beaucoup d'étiologies correspondent à des éléments exogènes (virus, bactéries,...). Afin de répondre au second cas d'utilisation vu en introduction (permettre au médecin de trouver ra-

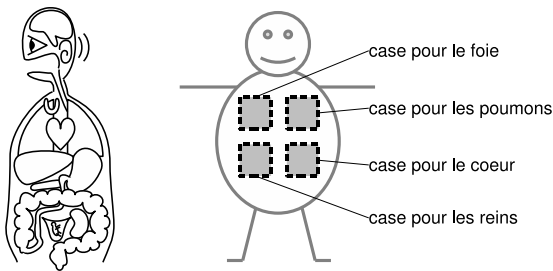


FIG. 4 – Exemples de schémas anatomiques. Le schéma de gauche est réaliste, tandis que celui de droite est simplifié et découpé en cases.

pidement l'absence d'une contre-indication, interaction médicamenteuse ou effet indésirable) nous avons présenté l'absence d'item pour une localisation anatomofonctionnelle ou une étiologie de manière explicite sur le contexte, en grisant la case correspondante.

Un prototype a été évalué sur un groupe de 11 médecins généralistes formés à l'usage de VCM par l'utilisation d'un didacticiel qui leur a été remis un mois à l'avance, et auquel ils ont consacré en moyenne 4h. L'évaluation avait pour objectif de comparer la vitesse de lecture et le nombre d'erreurs obtenus par des médecins utilisant tantôt une interface textuelle standard, tantôt l'interface graphique. Elle a porté sur des RCP chimériques générés de manière aléatoire à partir de vrais RCP issus de la base médicamenteuse Thériaque, et des questions de la forme "Ce médicament peut-il être prescrit sans précaution particulière chez un patient souffrant de la pathologie X ?", "Ce médicament peut-il être prescrit sans précaution particulière chez un patient prenant le médicament Y ?", ou "Ce médicament peut-il provoquer l'effet indésirable Z ?", tirées au hasard elles-aussi et correspondant au second cas d'utilisation donné en introduction.

3.2 Résultats

Le schéma anatomique que nous avons conçu a la forme d'un bonhomme stylisé que nous avons appelé "Monsieur VCM". La figure 5 montre les différentes localisations et étiologies, chacune étant représentée par le pictogramme correspondant en langage VCM. Seuls la tête, les pensées, le corps et un bras du bonhomme sont représentés, délimitant ainsi 5 zones : la tête (avec des localisations comme les yeux, les oreilles,...), les pensées (psychiatrie, psychologie,...), le corps (système digestif, coeur, organes sexuels,...), le bras (système nerveux périphérique, os, peau,...) et un espace en dehors du bonhomme où sont repré-

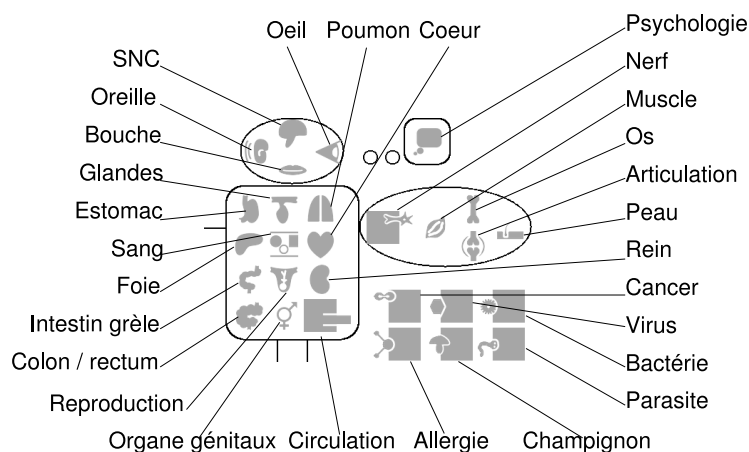


FIG. 5 – La répartition des différentes localisations et étiologies sur "Monsieur VCM".

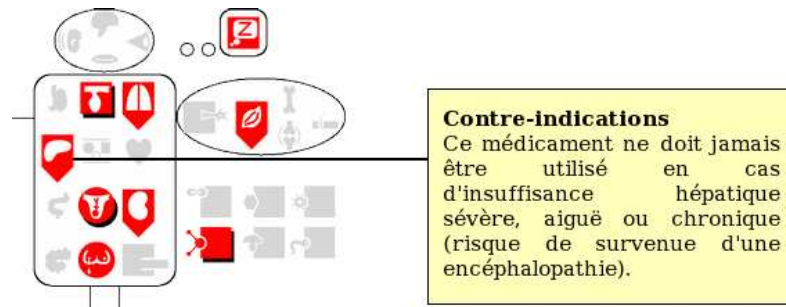


FIG. 6 – Exemple d'interface “Monsieur VCM” présentant les contre-indications du Stilnox[®] (un somnifère), après que l'utilisateur ait cliqué sur l'icône “insuffisance hépatique” placée sur la case du foie.

sentées les étiologies (virale, bactérienne,...). Les jambes et le second bras sont ébauchés, afin de compléter le dessin du bonhomme.

“Monsieur VCM” peut accueillir les icônes VCM correspondant aux contre-indications, aux effets indésirables ou aux interactions médicamenteuses (figure 6, un “Monsieur VCM” pour chacune des trois sections du RCP). Lorsque plusieurs icônes devraient occuper la même case du schéma, les relations *est-un* existant entre les icônes VCM sont utilisées pour les remplacer par l'icône du plus petit parent commun. Par exemple les icônes pour “insuffisance cardiaque” et pour “angor” se placent toutes deux au niveau du coeur ; elles seront remplacées par l'icône plus générale “pathologie cardiaque”. Une ombre est ajoutée à cette icône, pour indiquer qu'elle résulte d'une combinaison. Lorsqu'une contre-indication, un effet indésirable ou un médicament est associé à la fois à une localisation et une étiologie, son icône est présente deux fois. Au contraire, si une pathologie ou un médicament n'a ni localisation ni étiologie, son icône est placée en dessous du bonhomme ; ce cas reste cependant exceptionnel, comme nous l'avons montré lors de l'analyse des RCP. Lorsqu'une case ne reçoit aucune icône, celle-ci est occupée par un pictogramme grisé. Enfin, il suffit de cliquer sur une icône pour faire apparaître le ou les passages de texte correspondant dans le RCP.

L'accès à un RCP dans l'interface graphique se fait par une recherche portant sur le nom du médicament. Les trois sections contre-indications, effets indésirables et interactions médicamenteuses du RCP peuvent être représentées par trois “Messieurs VCM”. Il est aussi possible de regrouper les contre-indications et les interactions médicamenteuses sur le même “Monsieur VCM”, car le langage VCM définit des relations *est-un* entre les icônes de pathologies et de médicaments. En effet, ces icônes représentent en fait, du point de vue d'un médecin, un patient souffrant d'une pathologie et un patient prenant un médicament. Comme les médicaments sont représentés à l'aide de la pathologie qu'il traite, nous avons bien la relation *est-un* suivante : un patient prenant un médicament traitant la pathologie X est un patient souffrant de la pathologie X.

Les résultats de l'évaluation ont montré que l'interface utilisant “Monsieur

VCM” avait été lue par les médecins en moyenne 2,2 fois plus vite ($p < 10^{-16}$, test t apparié) et avec 1,7 fois moins d’erreurs ($p = 0.034$, test de Fisher exact, 16 erreurs contre 27), par rapport à l’interface textuelle similaire à celles existantes (Lamy *et al.*, 2008c). Le temps de réponse avec l’interface textuelle augmente fortement avec la longueur du RCP, et lorsque l’on demande au médecin de rechercher un item qui est absent du RCP. Avec l’interface graphique, la longueur du RCP a une influence plus faible sur le temps de réponse, et la recherche d’un item présent ou absent n’a aucune influence, ce qui était attendu car “Monsieur VCM” représente de manière explicite l’absence d’item concernant une localisation anatomo-fonctionnelle ou une étiologie.

4 Discussion et conclusion

4.1 Analyse d’un corpus de texte

L’analyse d’un corpus de RCP a permis de sélectionner les deux axes principaux de l’interface graphique. Le corpus comprenait 258 RCP sur un total d’environ 5000, cependant ces RCP sont représentatifs de l’ensemble et, faisant partie de la liste des médicaments essentiels de l’OMS, sont vraisemblablement parmi les plus utilisés. Nous avons réalisé une analyse assez simple ; des méthodes plus sophistiquées auraient pu être mise en oeuvre mais n’était pas nécessaires. Une approche similaire a déjà été utilisée pour analyser la section indication du RCP (Duclos & Venot, 2000). Les étapes manuelles ont été restreintes à l’ensemble des mots présents dans les candidats termes (sans les doublons), et non à l’ensemble de tous les candidats termes. Nous n’avons fait appel qu’à un seul expert, car la variabilité inter-expert est vraisemblablement limitée vu la faible granularité de l’étude ; il pourrait être intéressant de vérifier ce point. Les listes de mots marqués par l’expert pourrait resservir à d’autres études de ce type. Enfin, l’utilisation d’un filtrage aurait pu permettre de retirer les candidats termes non interprétables avant leur étude.

Les résultats de l’analyse montrent l’importance de la dimension anatomo-fonctionnelle. Ce résultat n’est pas surprenant car on retrouve fréquemment cette dimension, par exemple dans les classifications médicales. En revanche, il est intéressant de constater que l’utilisation conjointe des deux attributs localisation anatomo-fonctionnelle et étiologie permet de couvrir la quasi-totalité des contre-indications et des effets indésirables. Pour les médicaments, la plupart sont associés à une pathologie traitée qui peut être caractérisée par au moins un de ces deux attributs, mais il reste cependant des classes chimiques ou pharmacologiques qui ne correspondent pas directement à une pathologie traitée, ainsi que certains médicaments qui peuvent traiter plusieurs pathologies.

4.2 Interface graphique “Monsieur VCM”

“Monsieur VCM” répond aux besoins définis en introduction, en offrant une vue d’ensemble graphique des contre-indications, effets indésirables ou interac-

tions médicamenteuses d'un RCP, et en indiquant de manière explicite, par un pictogramme grisé, l'absence d'élément d'information concernant une localisation anatomo-fonctionnelle ou une étiologie. Les mêmes cases étant toujours placées au même endroit, "Monsieur VCM" permet de trouver rapidement les informations correspondant à une localisation ou une étiologie donnée.

L'une des particularités de cette interface est de combiner ensemble les icônes qui se rapportent à la même localisation anatomo-fonctionnelle ou à la même étiologie, en créant une icône plus générale. Ce mécanisme fait varier le niveau d'abstraction des icônes en fonction de la quantité d'information disponible, de sorte à avoir un "Monsieur VCM" dont la taille reste constante : un RCP plus long se traduira par un "Monsieur VCM" avec des icônes plus générales mais pas plus volumineux. Une autre particularité est de représenter de manière explicite l'absence d'éléments concernant une localisation ou une étiologie donnée.

Plusieurs auteurs ont proposé des interfaces employant des schémas anatomiques (Kirby *et al.*, 1996; McCullagh *et al.*, 2003; Sundvall *et al.*, 2007). Comparés à ces interfaces, "Monsieur VCM" a l'avantage d'utiliser un langage iconique, et de ne pas être purement anatomique grâce à l'ajout d'un axe étiologique. Par ailleurs, les interfaces précédemment publiées étaient destinées à la visualisation de données médicales et non de connaissances, et proposaient des schémas anatomiques plus réalistes. Pour la visualisation des connaissances, un schéma moins réaliste permet un plus haut niveau d'abstraction ; par exemple, "Monsieur VCM" ne distingue pas le rein gauche et le rein droit car il n'existe pas de contre-indication spécifique à l'un des reins. En revanche, pour visualiser les données médicales d'un patient, il peut être utile de distinguer les deux. Enfin, la généralisation de l'interface graphique iconique à d'autres cultures que la culture occidentale reste à évaluer.

4.3 Conclusion

Nous avons présenté une méthode pour concevoir une interface graphique pour l'accès aux contre-indications, effets indésirables et interactions médicamenteuses. La méthode repose sur l'analyse d'un corpus de texte afin de dégager les principaux axes permettant d'organiser les pathologies et les médicaments, puis sur l'utilisation conjointe de techniques de visualisation et d'un langage graphique. L'évaluation de cette interface auprès d'un groupe de médecins a donné de bons résultats.

Nous envisageons à présent la conception d'interfaces pour d'autres types de connaissances médicales, comme celles contenues dans les guides de bonnes pratiques cliniques. Une autre perspective est la construction d'une ontologie pour les icônes VCM et les "Messieurs VCM", afin d'automatiser certains processus comme la correspondance entre les icônes et les classifications médicales existantes. Par ailleurs, les problèmes de volume de connaissances ne sont pas spécifiques au domaine médical, et nous pensons qu'une approche similaire pourrait être appliquée dans d'autres domaines, par exemple à des textes de lois.

Références

- ANDREWS K. (2002). *Information visualisation : tutorial notes*. Graz University of Technology.
- BOUILLON P. & VANDOOREN F. (1998). *Traitement automatique des langues naturelles*. Paris, Bruxelles : De Boeck Université.
- BOURIGAULT D. (1995). LEXTER, a terminology extraction software for knowledge acquisition from texts. In *9th knowledge acquisition for knowledge based system workshop (KAW '95)*, Banff, Canada.
- CHANG S. (1986). *Visual Languages*, chapter Iconic Visual Languages, p. 1–7. Plenum Press, New York.
- DUCLOS C. & VENOT A. (2000). Structured representation of drug indications : lexical and semantic analysis and object-oriented modeling. *Methods Inf Med*, **39**, 83–87.
- FURNAS G. (1986). Generalized Fisheye views. In *Proceedings of the Human Factors in Computing Systems CHI '86 conference*, p. 16–23.
- HASCOËT M. & BEAUDOUIN-LAFON M. (2001). Visualisation interactive d'information. *Information, interaction, intelligence (I3)*, **1**(1), 77–108.
- KIRBY J., COPE N., SOUZA A., FOWLER H. & GAIN R. (1996). The PEN&PAD data entry system : from prototype to practical system. In J. BRENDER, J. CHRISTENSEN, J.-R. SCHERRER & P. MCNAIR, Eds., *Medical Informatics Europe (MIE-96)*, p. 430–434, Copenhagen : IOS Press.
- LAMY J.-B., DUCLOS C., BAR-HEN A., OUVARD P. & VENOT A. (2008a). An iconic language for the graphical representation of medical concepts. *BMC Medical Informatics and Decision Making*, **8**(16).
- LAMY J.-B., DUCLOS C., RIALLE V. & VENOT A. (2008b). Quelle méthodologie tenant compte des sciences cognitives pour la conception des langages graphiques ? *numéro spécial de la revue d'intelligence artificielle (RIA)*, **22**(3-4), 265–280.
- LAMY J.-B., VENOT A., BAR-HEN A., OUVARD P. & DUCLOS C. (2008c). Design of a graphical and interactive interface for facilitating access to drug contraindications, cautions for use, interactions and adverse effects. *BMC Medical Informatics and Decision Making*, **8**(21).
- MCCULLAGH P., MCGUIGAN J., FEGAN M. & LOWE-STRONG A. (2003). Structure data entry using graphical input : recording symptoms for multiple sclerosis. *Stud Health Technol Inform*, **95**, 673–8.
- MEUNIER J.-G. (1998). The categorial structure of iconic languages. *Theory & Psychology*, **8**(6), 805–825.
- MOORE T., COHEN M. & FURBERG C. (2007). Serious Adverse Drug Events Reported to the Food and Drug Administration, 1998-2005. *Arch Intern Med*, **167**, 1752–1759.
- PAIVIO A. (1990). *Mental representations : a dual coding approach*. Oxford University Press, Incorporated.
- SUNDVALL E., NYSTRÖM N., FORSS M., CHEN R., PETERSSON H. & ÅHLFELDT H. (2007). Graphical overview and navigation of Electronic Health Records in a prototyping environment using Google Earth and openEHR Archetypes. In K. KUHN, J. WARREN & T.-Y. LEONG, Eds., *MEDINFO 2007*, p. 1043–1047, Brisbane, Australia : IOS Press.